Session I: Finite Difference Solutions of Differential Equations, Part 2

I.1 Introduction

In the last session we ended up looking at the general solution of homogeneous, second order linear differential equations. We showed how a trial solution of the form $y_j = \beta^j$ can be used to obtain the full solution. It is sometimes advantageous to work with a different trial solution of the form $y_j = e^{mj}$. Suppose that we have a second order differential equation that can be represented in finite difference form as:

$$y_{j+2} - by_{j+1} + y_j = 0 \tag{I.1}$$

The coefficient, b, will depend on the particular equation we are trying to solve. If you are unclear about this, look back at the last session before continuing.

If we substitute our trial solution $y_j = e^{mj}$, we obtain:

$$e^{m(j+1)}(e^m - b + e^{-m}) = 0$$

which can be solved to find the values of m. Provided that $b > 2$ we can write[1].

$$2\cosh(m) = b$$

or

$$m = \cosh^{-1}(b/2) = \pm\alpha$$

Thus the full solution is:

$$y_j = c_1 e^{\alpha j} + c_2 e^{-\alpha j}$$

Note that the full solution can also be written in terms of cosh and sinh:

$$y_j = c_3\cosh(\alpha j) + c_4\sinh(\alpha j)$$

I.2 The heat flow equation

Having shown the general form of the full solution with this new trial solution, we shall use this alternative form of the solution to study a detailed example: the heat flow equation:

$$\frac{\partial T(x,t)}{\partial t} = \frac{\partial^2 T(x,t)}{\partial x^2} \tag{I.2}$$

Here, T = temperature, t = time, x = position variable. If you look up this equation in textbooks, note that the heat transport coefficient has been normalised to 1. Note also that this problem is in two independent variables, position and time.

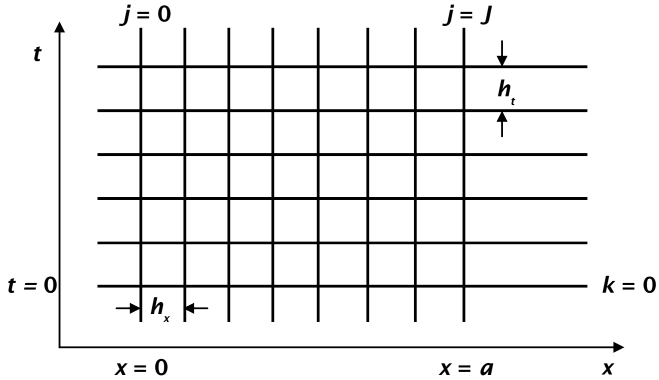
The first step to finding a numerical solution is to identify a suitable difference form of equation I.2. In this case, I've chosen to use the forward difference form for the first order differential term with

[1] Remember that $\sinh x = (e^x - e^{-x})/2$ and $\cosh x = (e^x + e^{-x})/2$ and that $\cosh x \geq 1$

respect to time and the central difference form for the second order term with respect to x, such that:

$$\frac{\Delta T(x,t)}{h_t} = \frac{\delta^2 T(x,t)}{h_x^2} \tag{I.3}$$

We require the solution in the semi-infinite strip; $0 \le x \le a$, $t \ge 0$.:



We shall apply the equation at the points given by the mesh (j, k) where:

1. $j = 0$ corresponds to $x = 0$
2. $j = J$ corresponds to $x = a$
3. $k = 0$ corresponds to $t = 0$.

Thus, with a mesh step size of h_x in the x direction, $a = Jh_x$.

We shall apply the boundary conditions

$$T_{0,k} = T_{J,k} = 0 \qquad \text{for all } t$$

That is, for all values of time, k, the temperature is zero at $j = 0$ and $j = J$.

In terms of the spatial variable, x, we are saying that the temperature is zero at all times at the locations given by $x = 0$ and $x = a$. We shall also apply a known temperature spatial distribution at time zero, which we can describe by the function f_j so that:

$$T_{j,0} = f_j \qquad 0 \le j \le J$$

I.2.1 The solution

Suppose the function $T_{j,k}$ is separable, so we can write $T_{j,k} = R_j S_k$. In other words, the temperature is defined by two functions R and S, where R is a function of space only, and S is a function of time only.

Substituting this into equation I.3 and dividing through by $T_{j,k}$ ($=R_j S_k$) we obtain:

$$\frac{\Delta S_k}{h_t S_k} = \frac{\delta^2 R_j}{h_x^2 R_j}$$

(This is very similar to the method of solving the diffusion of neutrons in the graphite stack experiment that you will meet in the laboratory.)

Now, because each part is only a function of either time or space, the only way they can be equal is if both are equal to a constant:

$$\frac{\Delta S_k}{h_t S_k} = \frac{\delta^2 R_j}{h_x^2 R_j} = -\alpha^2$$

We can now consider the two resulting equations:

$$\begin{aligned} S_{k+1} - S_k + \alpha^2 h_t S_k &= 0 \\ R_{j+1} - 2R_j + R_{j-1} + \alpha^2 h_x^2 R_j &= 0 \end{aligned}$$

Applying our boundary conditions:

$$R_0 = R_J = 0, \qquad R_j S_0 = f_j$$

Both these equations can be solved by the methods outlined above. In fact, the time part, S, can be solved, by simply considering the recursion relationship:

$$S_{k+1} = S_k \left(1 - \alpha^2 h_t\right)$$

from which we can see that:

$$\begin{aligned} S_1 &= S_0 \left(1 - \alpha^2 h_t\right) \\ S_2 &= S_1 \left(1 - \alpha^2 h_t\right) = S_0 \left(1 - \alpha^2 h_t\right)^2 \qquad \vdots \\ S_k &= S_0 \exp\left[\ln\left(1 - \alpha^2 h_t\right)^k\right] \\ &\text{or} \\ S_k &= S_0 \exp\left[k \ln\left(1 - \alpha^2 h_t\right)\right] \end{aligned}$$

We now have to solve for the spatial part. Here it is again:

$$R_{j+1} - \left(2 - \alpha^2 h_x^2\right) R_j + R_{j-1} = 0$$

If we substitute R_j with the alternative trial solution:

$$R_j = e^{mj}$$

then

$$e^{mj}\left(e^m - \left(2 - \alpha^2 h_x^2\right) + e^{-m}\right) = 0$$

Thus, other than the trivial solution $e^{mj} = 0$, we have our solution from:

$$e^m + e^{-m} = \left(2 - \alpha^2 h_x^2\right)$$

Since $\left(2 - \alpha^2 h_x^2\right) \le 2$, we can put $m = i\theta$, (where $i = \sqrt{-1}$), and write[2]:

$$e^{i\theta} + e^{-i\theta} = 2\cos(\theta) = \left(2 - \alpha^2 h_x^2\right)$$

[2] Remember that $\cos x = (e^{ix} + e^{-ix})/2$.

or

$$\alpha^2 = \frac{2}{h_x^2}\left[1 - \cos(\theta)\right]$$

So our solution $R_j = e^{i\theta j}$ can be written directly as:

$$R_j = A\cos(\theta j) + B\sin(\theta j)$$

where A and B are arbitrary constants. We can identify these constants from our boundary conditions. Firstly, since at $j = 0$ and $j = J$, R (the spatial part) must be zero, it follows that:

$$R_0 = R_J = 0$$

Thus for R_0:

$$R_0 = A\cos(\theta.0) + B\sin(\theta.0) = 0 \quad \text{i.e. } A \text{ must} = 0$$

Secondly, from the condition that R is zero at $j = J$, (and substituting $A = 0$):

$$R_J = B\sin(\theta J) = 0 \quad \text{i.e. } \sin(\theta J) \text{ must} = 0.$$

Therefore, there are conditions on the product θJ, namely $\theta J = 0, \pi, 2\pi, \ldots, n\pi$. Hence, $\theta = n\pi/J$ with $n = 0, 1, 2, \ldots$. It follows that:

$$\begin{aligned}\alpha^2 &= \frac{2}{h_x^2}\left[1 - \cos\left(\frac{n\pi}{J}\right)\right], \qquad n = 0, 1, 2, \ldots \\ &\equiv \alpha_n^2\end{aligned}$$

For each value of n there will be the corresponding constant B_n. So, our complete solution is:

$$T_{j,k} = R_j S_k = \sum_{n=1}^{J-1} a_n \sin\left(\frac{n\pi j}{J}\right) \exp\left[k \ln\left(1 - \alpha_n^2 h_t\right)\right] \tag{I.4}$$

where we have written a_n for the product of the constant B_n and the initial value S_0 of the time dependent function. Note that the contribution at $n = J$ is zero and that the solution repeats cyclically for $n > J$. Therefore the summation is only to $J - 1$. The constants, a_n, can be determined from our initial condition $T_{j,0} = R_j S_0 = f_j$ (the temperature distribution at $t = 0$):

$$T_{j,0} = R_j S_0 = \sum_{n=1}^{J-1} a_n \sin\left(\frac{n\pi j}{J}\right) \exp\left[0 \cdot \ln\left(1 - \alpha_n^2 h_t\right)\right] = f_j$$

We can use the orthogonality property of sinusoidal functions:

$$\sum_{n=1}^{J-1} \sin\left(\frac{n\pi j}{J}\right) \sin\left(\frac{m\pi j}{J}\right) = \frac{J}{2}\left(1 - \delta_{n0}\right)\delta_{nm}$$

where δ_{nm} is the Kronecker delta ($\delta_{nm} = 1$ if $n = m$, $\delta_{nm} = 0$ otherwise).

So, if we put $k = 0$ in our solution, and multiply by $\sin(m\pi j/J)$ and sum over all j, we obtain:

$$a_m = \frac{2}{J}\sum_{j=1}^{J-1} f_k \sin\left(\frac{m\pi j}{J}\right)$$

which allows us to isolate each coefficient and hence complete our solution.

We should examine the effect of arbitrarily decreasing the size of the meshes h_x and h_t. Such a procedure would seem reasonable if we wished to increase the accuracy of the calculation. Consider the following expression in the limit of vanishing mesh size:

$$\lim_{h_t \to 0} \left[k \ln \left(1 - \alpha_n^2 h_t \right) \right] \tag{I.5}$$

By expanding the log term, we see that:

$$\lim_{h_t \to 0} \left[k \left(-\alpha_n^2 h_t - \frac{1}{2!} [\alpha_n^2 h_t]^2 - \frac{1}{3!} [\alpha_n^2 h_t]^3 \ldots \right) \right]$$

Putting $t = k h_t$, this limit actually becomes $-\left(\frac{n\pi}{a}\right)^2 t$ which, when we take the exponential, gives for the time variation portion of equation I.4, $\exp\left[-\left(\frac{n\pi}{a}\right)^2 t \right]$. This is precisely the solution that we would have obtained if we had solved the heat transport equation analytically.

Exercise I.1

1. Make a spreadsheet for the following conditions:

$$\begin{aligned} a &= 10 \\ h_x &= 0.1 \\ \text{Hence} \quad J &= 100 \\ \text{Initially set:} \quad n &= 2 \\ \text{Hence} \quad \alpha_n^2 &= 0.394654 \end{aligned}$$

Starting with $h_t = 0.001$, make a table for different times and the corresponding values of k. (It will help later calculations if you put h_t in a location in the sheet and use an absolute address reference to it when calculating the k column). Setup your worksheet as shown in the example below:

t	k	A	B	C	D
0	0				
1	1000				
2	2000				
3	3000				
...	...				
100	100000				

In column A evaluate $\left[k \left(-\alpha_n^2 h_t \right) \right]$ and in column B $\left[k \left(-\frac{1}{2!} [\alpha_n^2 h_t]^2 \right) \right]$, thus showing that the terms in the expansion of equation I.5 are getting rapidly smaller.

In column C evaluate exp(A+B).

Finally in column D evaluate $\exp\left[-\left(\frac{n\pi}{a}\right)^2 t \right]$ and hence show that the limit applied to equation I.5 gives a time function (column C) that is indeed leading to the analytic answer. However, there is a problem with what we have done.

Exercise I.2

1. Redo your spreadsheet with $h_t = 0.1$. Is the value in column C still a good approximation to column D?
2. Try other sizes for h_x and h_t and see how the approximate solution compares with the analytic solution.

The reason we can have problems is that in using equation I.5 we have taken the limit without considering too much about what has been happening to α_n^2.

Equation I.5 can also be written in the following way:

$$\lim_{h_t \to 0} \left[k \ln \left(1 - \alpha_n^2 h_t\right)\right] = \lim_{h_t \to 0} \left[k \ln \left(1 - \frac{2h_t}{h_x^2} \left\{1 - \cos\left(\frac{n\pi}{J}\right)\right\}\right)\right]$$

As the result of decreasing the mesh size, J becomes much larger. Moreover, for values of n a little less than and about J, $\cos(n\pi/J) \approx -1$. So:

$$1 - \frac{2h_t}{h_x^2} \left\{1 - \cos\left(\frac{n\pi}{J}\right)\right\} \approx 1 - \frac{4h_t}{h_x^2}$$

We must now consider:

$$\lim_{h_t \to 0} \left[k \ln \left(1 - \frac{4h_t}{h_x^2}\right)\right]$$

Obviously this limit only exists if $h_t/h_x^2 \leq 1/4$. Therefore when we are decreasing our mesh sizes, h_t and h_x, we must do so in such a way that $h_t/h_x^2 < 1/4$ otherwise we will obtain a divergent solution rather than converging on the correct one. This is a good example of how some care needs to be exercised in the use of a computer code.

I.3 Parasitic solution

There is a second problem to be aware of when seeking a numerical solution to differential equations. Let us consider a more straightforward example:

$$\frac{dy}{dx} = -\alpha y(x)$$

We can write this in simple forward-difference form as:

$$y_{j+1} - y_j = -\alpha h y_j$$

which (compare with Example 2 in Session H) has the solution:

$$y_j = y_0 \exp\left[j \ln(1 - \alpha h)\right]$$

Expanding the log term:

$$\ln(1-\alpha h) = -\alpha h - \frac{(\alpha h)^2}{2} - \frac{(\alpha h)^3}{3} \dots$$

and ignoring terms $> h^2$, we obtain:

$$y_j = y_0 \exp\left(-j\alpha h - j\frac{(\alpha h)^2}{2}\right)$$

Putting $x_j = jh$:

$$y(x_j) = y_0 \exp(-\alpha x_j) \exp\left(-j\frac{(\alpha h)^2}{2}\right)$$

Since the analytic solution of this simple first order differential equation is simply:

$$y(x_j) = y_0 \exp(-\alpha x_j)$$

we have an error term of the order of $\exp\left(-j\frac{(\alpha h)^2}{2}\right)$. As $h \to 0$, $y \to y_{analytic}$.

Suppose we now try to improve the accuracy of our method by using a different finite difference operator:

$$\frac{dy}{dx} = \frac{y_{j+1} - y_{j-1}}{2h} + O(h^2)$$

This is of higher truncation order $[O(h^2)]$ than the forward difference operator, which as we have seen is of order $O(h)$.

Our equation now becomes:

$$y_{j+1} - y_{j-1} = -2\alpha h y_j$$

The solutions are:

$$y_j = c_1\left(\sqrt{1+\alpha^2 h^2} - \alpha h\right)^j + c_2\left(-\sqrt{1+\alpha^2 h^2} - \alpha h\right)^j$$

where c_1 and c_2 are arbitrary constants. Expanding the square root we obtain:

$$\sqrt{1+\alpha^2 h^2} = 1 + \frac{1}{2}\alpha^2 h^2 - \frac{1}{8}\alpha^4 h^4 + \dots$$

Ignoring terms beyond h^4, the first part of the solution, $y_1(j)$, becomes:

$$\begin{aligned} y_1(j) &= \left(1 - \alpha h + \frac{1}{2}\alpha^2 h^2 - \frac{1}{8}\alpha^4 h^4\right)^j \\ &= \exp\left[j\left(-\alpha h + \frac{1}{6}\alpha^3 h^3\right)\right] \quad \text{to the same order in } h \end{aligned}$$

So:

$$y_1(j) = y_0 \exp(-\alpha x_j) \exp\left(j\frac{(\alpha h)^3}{6}\right)$$

This certainly converges faster than the case when we used the forward difference operator. However, consider the second part of the solution:

$$\begin{aligned} y_2(j) &= \left(-\sqrt{1+\alpha^2 h^2} - \alpha h\right)^j \\ &= (-1)^j\left(\sqrt{1+\alpha^2 h^2} + \alpha h\right)^j \end{aligned}$$

For all $h > 0$ this function is ever increasing in magnitude and oscillatory, and will eventually dominate the solution. By writing the solution down in closed form, as we have here, we can simply put $c_2 = 0$ to rid us of this term. If we had attempted to solve our question differently by, say, matrix inversion, this would have produced y_j directly and not allowed us to put $c_2 = 0$. The presence of this parasitic solution would be problematic. A parasitic solution must occur in this case because, with our second operator, we have a 2^{nd} order finite difference equation, which demands two solutions. Considerable care must be taken in these cases to ensure that the proper convergence of these extra solutions does exist.

In this section we have only considered homogeneous equations. Inhomogeneous equations can be similarly considered but in practice these usually require matrix inversion techniques for obtaining the particular integral. Consequently they are usually better treated by conventional matrix methods that obtain the complete solution directly.